Methoden

Richard Matthaei* and Markus Maurer

Autonomous driving – a top-down-approach

Autonomes Fahren – ein Top-Down-Ansatz

Abstract: This paper presents a functional system architecture for an "autonomous vehicle" in the sense of a modular building block system. It is developed in a top-down approach based on the definition of the functional requirements for an autonomous vehicle and explicitly combines perception-based and localization-based approaches. Both the definition and the functional system architecture consider the aspects operating by the human being, mission accomplishment, map data, localization, environmental and self-perception as well as cooperation. The functional system architecture is developed in the context of the research project "Stadtpilot" at the Technische Universität Braunschweig.

Keywords: Autonomous driving, functional system architecture, localization, maps, V2X-communication, perception, mission accomplishment, cooperation.

Zusammenfassung: In diesem Artikel stellen wir eine funktionale Systemarchitektur für ein "autonom fahrendes Straßenfahrzeug" vor, die im Sinne eines modularen Baukastensystems entworfen ist. Sie wurde in einem Top-Down-Ansatz ausgehend von einer Definition des Funktionsumfangs eines "autonom fahrenden Straßenfahrzeugs" entwickelt und führt explizit wahrnehmungsbasierte und lokalisierungsbasierte Ansätze zusammen. Sowohl die Definition des Funktionsumfanges als auch die funktionale Systemarchitektur berücksichtigen die Aspekte Bedienung, Missionsumsetzung, Karten, Lokalisierung, Umfeld- und Selbstwahrnehmung sowie Kooperation. Die Ergebnisse basieren unter anderem auf Erkenntnissen aus dem Projekt "Stadtpilot" der Technischen Universität Braunschweig.

Schlüsselwörter: Autonomes Fahren, funktionale Systemarchitektur, Lokalisierung, Karten, V2X-Kommunikation, Wahrnehmung, Missionsumsetzung, Kooperation.

***Corresponding author: Richard Matthaei,** Technische Universität Braunschweig, e-mail: matthaei@ifr.ing.tu-bs.de
Markus Maurer: Technische Universität Braunschweig

DOI 10.1515/auto-2014-1136
Received July 21, 2014; accepted December 7, 2014

1 Introduction

Today, the vision of "autonomous driving" is widely discussed. The media are reporting about the progress in this field of research and promise an introduction to the market in the near future. Additionally, automotive companies are competing in their race to develop new technologies, and software companies are competing with the vehicle manufacturers.

The estimates of an introduction to the market diverge significantly. All options from 3 to 30 years or even "never" are mentioned. This high variance might be caused, amongst others, by a heterogeneous understanding of the functional requirements for an "autonomous" vehicle.

Hence, in a first step (Section 2) we propose a definition of the functional requirements for an autonomous on-road motor vehicle which follows the vision of an automated taxi and thus equals the definition of full automation, SAE-level 5, according to SAE international [48]. These requirements include the usage of map data and communication technologies, absolute and global localization, environmental and self-perception, mission accomplishment, and the integration of the human being as a passenger and as another traffic participant in the near surrounding of the autonomous vehicle. Based on these definitions and based on the state of research (Section 3) we introduce a systemically developed functional system architecture (Section 4) which covers the formulated functional requirements in a top-down approach (Section 5). In Section 6 the proposed functional system architecture is checked against the aforementioned requirements.

Our system architecture allows the discussion of existing approaches for autonomous driving and fully, highly or conditionally automated systems as well as today's assistance systems. Some examples are given in Sections 5 and 6. With the help of the proposed functional system architecture it is also possible to visualize further needs of research for the development of autonomous vehicles.

2 Functional requirements for an autonomous on-road motor vehicle in public road traffic

In this article an autonomous vehicle is understood as an automated taxi, similar to the description of Wachenfeld and Winner [54]. It is able to move "freely", because it is not constrained to rails, power supply lines or a bus bar, and it drives in public road traffic.

The operating of the autonomous vehicle by the human being is done on a very *intuitive* level. This means, neglecting a service mode, the vehicle is only instructed by the input of a mission. Typically, a mission for an on-road vehicle consists of a transportation task. People, goods or just the vehicle itself might be transported. In future systems also surveillance and other tasks might be relevant for autonomous vehicles.

In case of transporting human beings the mission has to be adaptable to the current needs of the passengers at any time. Such an adaptation might be caused by triggering an emergency stop (see Wachenfeld and Winner [54]) or by adding a stopover at a restaurant, the next bathroom or a hospital. Thus, the operating interface of an autonomous vehicle is similar to the well known interfaces of today's navigation systems and is consequently provided at the most *abstract* level from the system's point of view (see Section 4.5).

The appliance to the public road traffic increases the demands on an autonomous vehicle (in this case also automated vehicle) concerning both the environmental perception and the driving behavior. The urban environment in particular puts high demands on the environmental perception. It is necessary that the vehicle robustly detects and classifies the stationary elements (e.g. road course, signs, traffic lights) and the movable elements (e.g. traffic participants, human beings, animals). It is mandatory due to consistency reasons that human beings and technical systems use the same optical features for orientation as they share the same road environment (see e.g. Bar Hillel et al. [4], Huang et al. [23]).

In this case of mixed traffic (automated and manually driven vehicles) the locally defined road traffic regulations are of special interest [54]. They define a minimal amount of environmental elements (signs, road markings, traffic participants etc.) which have to be perceived and considered. Additionally, the regulations specify the behavior in defined situations [54, p. 6]. The basic components of the road traffic regulations are the mutual considerateness, a clear behavior pattern as well as communication and cooperation.

In addition to these pure functional requirements it is mandatory within the meaning of responsible acting that automated vehicles do not constitute any danger to their environment. Therefor the vehicle needs to be aware of its skills and abilities and has to act accordingly to its actual state. So, the estimation of the skills and abilities including the surveillance of hard- and software is another mandatory requirement (on-board diagnostics). Moreover, the vehicle has to be resistant against misuse and manipulation.

In summary, autonomous vehicles have to handle at least the following aspects derived from the aforementioned requirements:

1. Operating: The vehicle has to be instructed by a human being, e.g. defining the mission.
2. Mission accomplishment: The vehicle has to accomplish the desired mission. This includes the navigation task, the behavior generation and the control of the actuators.
3. Map data: Map data is required for route planning purposes in particular.
4. Localization: The vehicle needs to know its global pose for the usage of map data (e.g. navigation tasks) and communication purposes (e.g. vehicle-to-vehicle (V2V) or vehicle-to-infrastructure (V2I) communication).
5. Environmental perception: The vehicle has to perceive its local stationary and movable environment, including the dynamics of the movable elements.
6. Cooperation: The vehicle has to react to the intentions of other traffic participants (automated vehicles and human drivers) and it has to communicate its own intentions to the other traffic participants.
7. Safety: It must be ensured that the vehicle does not constitute any danger to its environment.
8. Self-perception: The vehicle needs to be aware of its current state (functional capabilities of its components, motion etc.).

The aspects of interior surveillance as well as aspects of safety and security concerning misuse and manipulation are not discussed within this article.

3 Related Work

3.1 The human being as an archetype

Assuming a mixed traffic with human drivers and automated vehicles, approaches aspiring to model human behavior seem basic because today's traffic system is established by human beings and organized based on their abilities.

There are at least two well known approaches of modeling the human behavior, that of Rasmussen [46] and that of Donges [11]. Rasmussen [46] focuses on the goal oriented behavior of a human being in general using a three-level model. Donges [11, 12, 13] also uses a three-level model, but he concentrates on the driving task. According to his model, the driving task can be decomposed into three hierarchical tasks as well: navigating, guiding and stabilizing. In a first step, the mission is processed on the highest level performing route-planning and navigation tasks. These steps require data about the relevant road network. On the next level, the resulting schedule and route are processed based on the local scene in the near surroundings of the vehicle. This level is characterized by an open-loop control. The selected maneuvers of the guiding module are then processed by the corresponding closed-loop controllers on the stabilization level.

A very similar model was presented by Michon [35] which was set into relation to the levels of Rasmussen [46] by Hale et al. [17] (see Table 1).

Table 1: Matrix of tasks according to Hale et al. [17].

Processing level / Task	Skill-based	Rule-based	Knowledge-based
Planning (Navigation)	Home/work travel	Choice between familiar routes	Navigating in strange town
Maneuver (Guidance)	Negotiating familiar junctions	Passing other cars	Controlling a car on icy roads
Control (Stabilization)	Road-holding round corners	Driving an unfamiliar car	Learner on first lesson

3.2 Autonomous driving

In the field of autonomous driving many system-architectures were published in the last two decades. The early approaches by Dickmanns [9], Dickmanns et al. [10] are followed up by Maurer [34], Pellkofer [44] and Siedersberger [49] and integrated into the real-time control system (RCS) by Albus [1].

Additionally, during the DARPA challenges several architectures had been published e.g. by Bacha et al. [2], Baker and Dolan [3], Bohren et al. [6], Hurdus et al. [24], Leonard et al. [28], Miller et al. [37], Montemerlo et al. [38], Rauskolb et al. [47]. All of them focused on the clear constraints of the DARPA challenge, which only correlate partially with real urban scenarios (see e.g. Bar Hillel et al. [4]). Most of them decided to use a localization-based approach as detailed map data was available. Among the leading teams only Leonard et al. [28] explicitly developed a perception-based approach.

The research activities on autonomous driving at the Technische Universität Braunschweig based on the participation at the DARPA Urban Challenge continued in the research project "Stadtpilot". Wille [55, p. 102] for example presented a bottom-up description of the developed system for the status of development of 2010. The functional system architecture of the autonomous vehicle "Leonie" mainly consisted of an input column, covering maps, localization, and perception in a rough description, and an output column which shows the structure of action planning and action execution in more detail. It provided two abstraction levels, one for perception and execution and another for situation assessment and action planning.

In a further state of development presented by Nothdurft [42, p. 72] the architecture consisted of a three-level design providing a strategical, tactical and operational level. Route-planning was done on the strategical level, V2X-information was associated to the tactical level as well as the decision unit. Sensor data in general (for localization and perception purposes) was associated to the operational level as well as a trajectory planner and the vehicle control. The main aspect of that architecture was the "context model" which covered and processed all relevant environmental information. Thus, that architecture mainly provided a big data-base with all relevant environmental information. It did not suggest any kind of structure for the processing of the environmental data. This is similar to many other system architectures (e.g. Bacha et al. [2], Kammel et al. [25], Leonard et al. [28], Montemerlo et al. [38]) providing a single "perception"-block with sub-modules for the processing of all environmental data but without a hierarchical structure.

3.3 Need of research

During our research on autonomous driving in urban environment within the project "Stadtpilot" (especially with the central "context model") it turned out that environment modeling is not just tracking other vehicles but is much more complex as soon as we assume that the vehicle does not only drive within a "shared space"[1] but has to consider various rules according to local traffic regulations based on lane-markings, traffic lights, and traffic signs as well as cooperative aspects etc.

Modularization and hierarchical structuring are commonly known mechanisms to manage this high complexity. Looking at the system architectures of the DARPA Urban Challenges, this strategy for system development is already established for the processing of the mission: the entire driving task is commonly subdivided into multiple (in most cases three) sub-tasks. In many projects the modularization equals the ideas of Donges [11]. But there is a lack of modularization and structuring the entire system concerning the processing of the incoming data from various sources (e.g. map data, the own on-board perception system or the perception system of other traffic participants or infrastructure received via V2V-communication) and concerning the integration of cooperative aspects. According to the models of Donges [11] and Rasmussen [46], each abstraction level of the mission accomplishment requires a particular abstraction level for the representation of environmental features.

Because of the increasing complexity of the resulting sub-tasks for the entire driving task when driving autonomously in *urban* scenarios, one vital question is, how far it is possible to process environmental data independently from the current mission of the vehicle. Or in other words, which steps are really part of the "perception" and which ones already belong to a function specific interpretation of the perceived data. A clear separation would help to develop an application-independent environmental perception also for advanced driver assistant systems (discussed in e.g. Holder et al. [21]). "Application-independent" means in this case that the environmental perception has to fulfill the highest demands of the subsequent applications in each case and is no more specialized in a certain application like in current adaptive cruise control (ACC) or blind spot assistant systems.

Due to the aforementioned increasing complexity it is no longer convenient to represent the processing steps required for a consistent environmental model just with one single "perception"-block. Among the introduced architectures only Miller et al. [37], Albus [1] and Maurer [34] explicitly consider a stepwise abstraction of the environmental data. But even these proposals do not consider cooperation aspects and do not make a systematic usage of map data transparent.

Especially the usage of map data is often considered only as an additional data input to the environmental sensors (except for Miller et al. [37] and Leonard et al. [28]). Probably, detailed map data mislead the developers to use it not only for an expansion of the field of view but also as a replacement for algorithms and sensors for an environmental perception. In some approaches for example the course of the road is not perceived and modeled at all but the problem of road detection is reduced to a map-relative localization problem (e.g. Montemerlo et al. [38], Wille [55] or Szczot et al. [50]). These approaches in general are inconvenient regarding safety aspects, because the up-to-dateness of map-data cannot be ensured due to an interruption of the observation of the environment between creation and usage of map data. Only a real-time perception system (either on-board or as part of e.g. a smart infrastructure) can ensure the up-to-dateness of the environmental data. Thus, a real-time perception system for the entire scene (stationary and movable environmental features) has to be part of the functional system architecture.

Our motivation for the development of an overall functional system architecture is to combine many aspects of published system architectures in order to find a functional system architecture, which covers all aspects of an autonomous vehicle as defined in Section 2 in one system description and thus can also be used as a kind of building block system for the development of less complex systems with fewer functional requirements.

4 Functional system architecture

4.1 Overall architecture

In the context of this architecture proposal (see Figure 1) the single vehicle is understood as a part of a superordinate system. The developed architecture combines a subset of elements of published architectures having an inner-city intersection assistant and autonomous driving in mind.

The main structure of the system architecture is a three-level design similar to the multi-level designs of Bonasso et al. [7], Donges [12], Maurer [34], Nothdurft [42] and Du et al. [14]. Du et al. [14] also introduced three lev-

1 see e.g. Hamilton-Baillie [18]

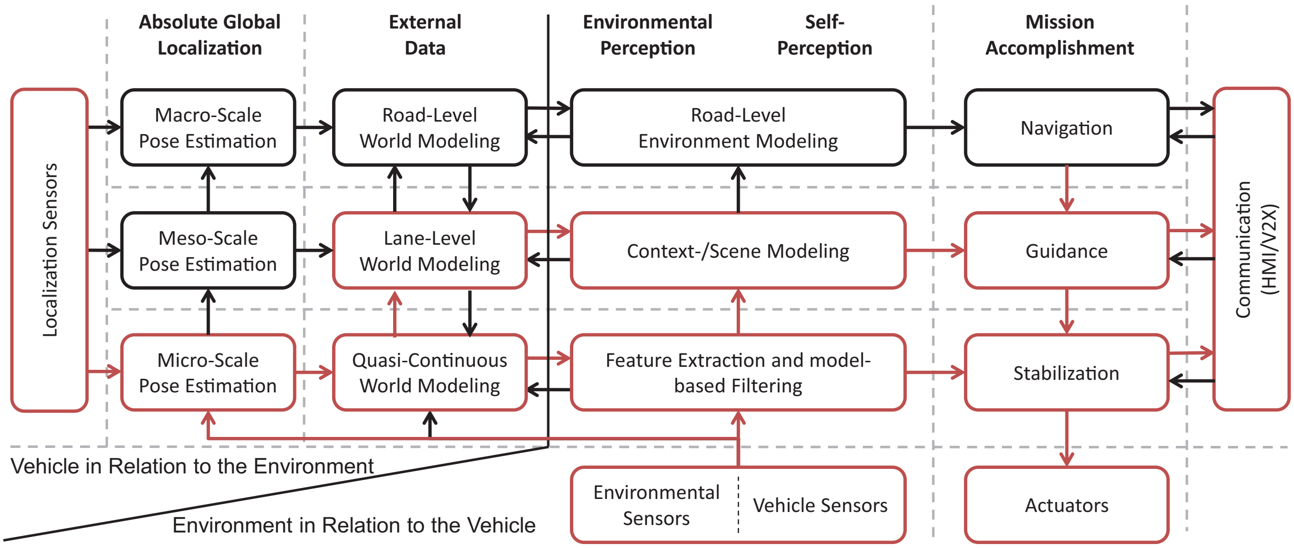

Figure 1: Functional system architecture for an autonomous on-road motor vehicle in the sense of a modular building block system. Modules currently developed in the research project Stadtpilot are marked in red. HMI: Human Machine Interface [30].

els of resolution which are assigned to the aforementioned three levels of the system architecture as follows:
– strategical level: planning, macro-scale resolution,
– tactical level: decision making, meso-scale resolution,
– operational level: reactive stabilization (micro-scale resolution).

These three levels differ (among other characteristics) in their resolution, horizon and accuracy (concerning time and space), relevant environmental features, tasks, and cycle times.

In an orthogonal direction to these three levels we introduce the columns “absolute global localization”, “external data”, “perception” (consisting of environmental perception and self-perception), and the “mission accomplishment”. This core of the system (consisting of vehicle and infrastructure) is framed by the sensors, actors and communication equipment for the exchange of data with human beings or other automated traffic participants.

The columns “perception” and “mission accomplishment” are state of research and already part of many system architectures (e.g. Baker and Dolan [3], Dickmanns [9], Leonard et al. [28], Montemerlo et al. [38], Wille [55]). They are typically part of a vehicle-referenced view which means that the environment is described *in relation to the vehicle*. An absolute global localization is not necessary in this case.

On the contrary, the absolute global localization and the external data describe the overall system “world and vehicle” from another perspective: They describe the environment in an absolute global reference frame while the global localization determines the pose of the vehicle *in relation to the environment*.

In the following the introduced columns “absolute global localization”, “external data”, “perception”, and “mission accomplishment” are discussed in detail.

4.2 Absolute global localization

The absolute global localization of a vehicle is required for two fields of application: the integration of external data (common usage of map data among multiple traffic participants and data exchange via V2V or V2I communication) and for the stabilization of the vehicle in environments without local environmental features (e.g. in deserts). The idea of automated map updates on a central server in particular (see Visintainer and Darin [53] and Deragården and Thomas [8]) can only be realized with the aid of an absolute global pose.

In today’s systems the incoming data is obtained from different global navigation satellite systems (GNSS). In the context of our architecture, the accuracy (trueness and precision) of the localization solution determines the applicable level. Today’s standard GNSS-receivers have positioning errors up to 20 m. Thus they are only applicable within the *strategical* level with a macro-scale pose. Certain methods allow to improve the position’s trueness (e.g. differential global positioning system, DGPS) and its precision (e.g. by fusing the motion estimation into the position estimation). These methods allow the appliance of the localization solution – as the case may be – at the *tactical* or even the *operational* level.

Depending on the safety requirements and the concrete system design, an absolute global pose with a micro-scale accuracy according to the restrictions of the operational level might be required. However, localization errors occur especially in urban environments due to the occlusion of satellites, multi-path propagation, or other effects. That is why state of the art systems do not guarantee a sufficient availability for autonomous driving (at least in urban environments) with a micro-scale accuracy of the pose [41, pp. 314f] and why Moore et al. [39] propose a local reference frame for the stabilization tasks without a direct GNSS-support.

4.3 External data

4.3.1 Processed data

All environmental data perceived or generated outside the host vehicle, e.g. by other vehicles or manually created, and provided via radio communication or data storage media, is part of the "external data". The "external data" may also be understood as a world model (not only environment model) and is typically defined within a global (or more general: common) reference frame. Otherwise it would not be applicable for multiple users. The module "external data" includes:

– data about the stationary environment, hereinafter referred to as scenery (map data, state of the traffic light, weather conditions, defined according to Geyer et al. [16]),
– data about the movable environment (traffic jam, hazard information, temporary road closures, V2X (V2I and V2V) object lists, etc.),
– requests and state information of other traffic participants via V2X,
– a global pose (map-relative or absolute) of the host vehicle which enables the usage of the V2X and map data.

Depending on the abstraction level within the architecture, external data provide different levels of accuracy, representations and time spans of change. For example, messages concerning traffic jams and hazards as well as temporal road closures are (depending on their level of detail) part of the *strategical* or *tactical* level. These messages are already applied in today's navigation systems and sent via traffic message channel (TMC) or internet to the vehicles.

The state of a traffic light with an association to a lane and thus a reduction of its position accuracy down to lane-level accuracy is also part of the representation of the *tactical* level (see Figure 2). An example of the *operational* level is given by object lists with hypotheses of traffic participants perceived by infrastructure (e.g. Homeier and Wolf [22]).

4.3.2 Input and output

As a first step, the architecture provides two data inputs into the pool of "external data": an absolute global pose and local environmental data in the sense of a cooperative or collaborative environmental perception. A third input is given by data derived from map suppliers. Additionally, the architecture also implicitly contains a fourth input: the reception of driver intentions via V2X. So, "external data" is designed as a central world model which is used simultaneously by all traffic participants.

4.3.3 Data processing

An elementary task within the processing of the module "external data" is the preparation of map data in such a way that it can be considered in the vehicle reference system for data processing within the module "environmental perception". This requires the determination of a *map-relative pose* which can be obtained by correspondent map-matching approaches (see e.g. Quddus et al. [45]). In a special case of a globally exact map, the resulting pose of the matching process equals the absolute global pose and thus leads to the process of a *map-aided localization* (currently not consistently used in literature).

Another task within the processing of external data is a fusion of the environmental features derived from multiple (traffic-) participants (world modeling).

On the *operational* level these features might be point landmarks (in an object-based representation) or parts of a grid-based representation (see also "view", "local map", "global map" in Elfes [15]). For on-road vehicles in structured environments the relevant features on the *tactical* level are described by the course of the lanes with associated speed limits or driving directions. All these features can be stored in a history, as well as the even more abstracted course of the road with its topological information on the *strategical* level.

For mapping purposes two different global localization solutions are required: a *map-relative pose* is required to update the semantical data (like traffic flow, traffic signs

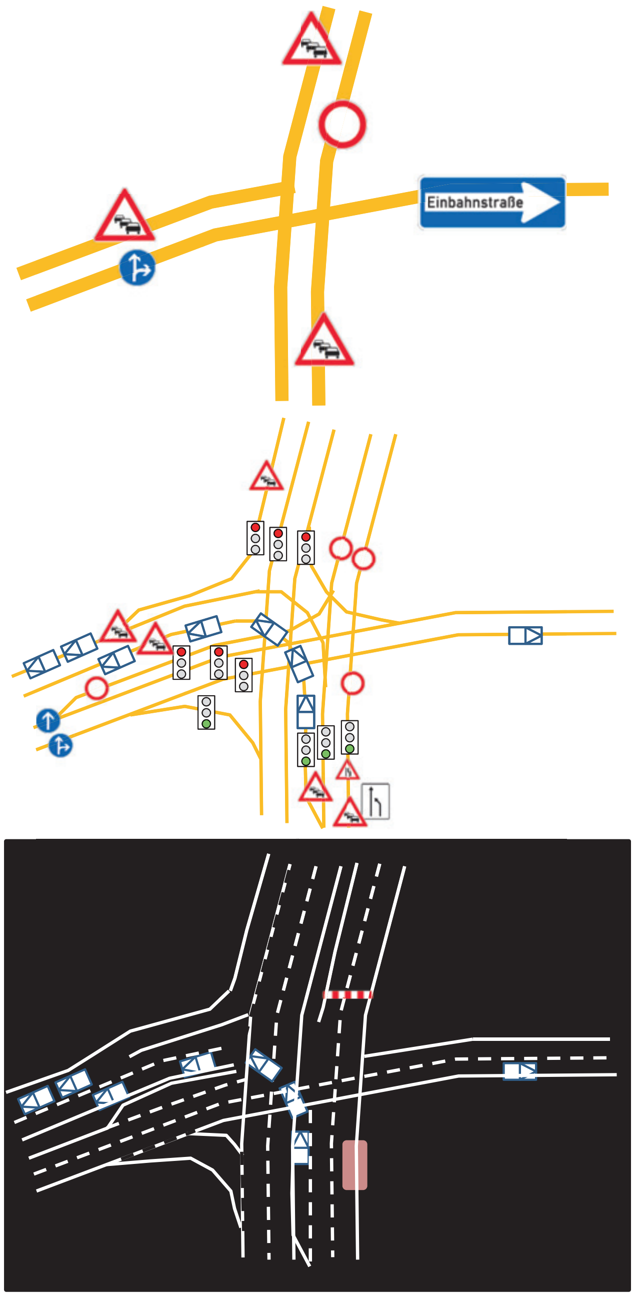

Figure 2: Illustration of the different representations of the vehicle's environment depending on the different abstraction levels, from top to bottom: strategical level, tactical level, operational level [30].

or number of lanes) and an *absolute global pose* is required in order to correct geometrical errors or add missing features like roads, lanes or point-landmarks (depending on the abstraction level).

In case an absolute global position is not part of the system, this mapping problem is reduced to the known simultaneous-localization-and-mapping-problem (SLAM-problem) (e.g. Thrun et al. [51]) and the loop-closure-problem (e.g. Milford and Wyeth [36]). However, an exchange of this locally stored data with other traffic participants is not possible due to the lack of a common reference frame. The vehicle then navigates in its *own* world.

Within map data a transition between the levels of abstraction is possible by introducing model assumptions, similar to the proceeding within the environmental perception (see Section 4.4). Take lane-level maps as an example: They can be derived from road-level maps (with a certain error) using the number of lanes and the lane width (e.g. see Knaup and Homeier [26] or Müller et al. [40]). This processing describes the reduction of the level of abstraction. To increase the level of abstraction the processing is done inversely: Lane-level maps can be derived from maps with detailed lane-borders and road-level maps can be generated from lane-level maps by calculating, e.g., a mean value of the lane's support points.

The different representations of the environment are illustrated in Figure 2.

4.4 Perception

According to the different representations of the environmental features within the module "external data", the environment is represented in different ways, depending on the level of abstraction within the environmental perception as well (see Figure 2).

4.4.1 Processed data

On the lowest, the *operational* level the focus lies on the extraction of precise and quasi-continuous values from incoming sensor data. The algorithms of the module "environmental perception" determine the size, position, velocity, color and other features of objects in the near surroundings of the vehicle. Thus mainly *geometrical* values are processed. The module "self-perception" processes data for the representation of the inner vehicle state.

The operational level's output to the higher tactical level covers, amongst others, the vehicle state, weather conditions, states of traffic lights and (variable message) signs, the course of the lane markings, the position and movement of other traffic participants, as well as an image of the stationary raised surroundings. All these environmental features are processed independently up to this step and also transmitted to the mission accomplishment for a high-frequent closed-loop control.

Typical algorithms of the operational level are methods for object and lane tracking or grid-based approaches

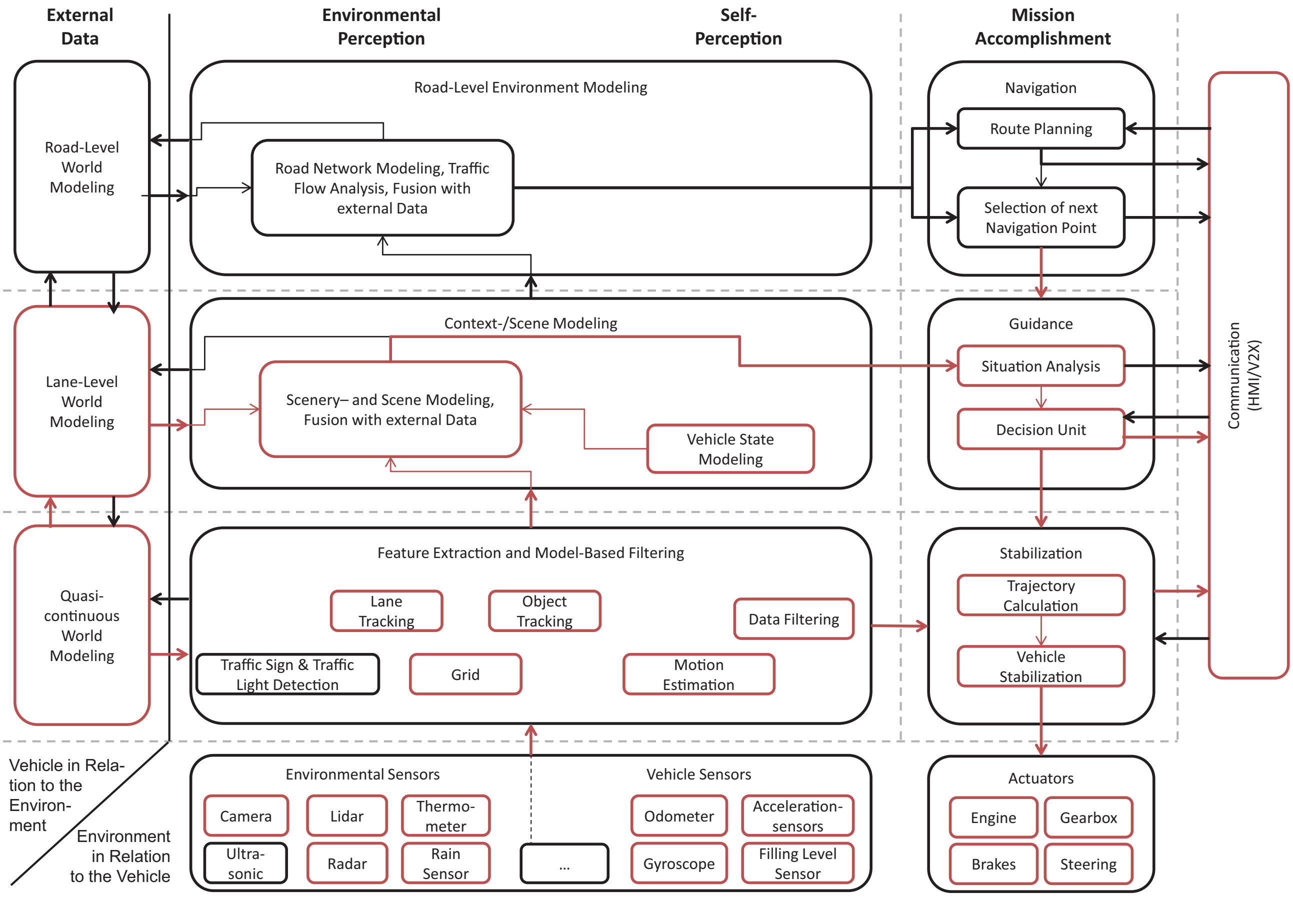

Figure 3: Assignment to the different abstraction levels of methods for the environmental perception, the self-perception, and the mission accomplishment. Sensors and actuators currently used and modules currently developed in the research project Stadtpilot are marked in red. For reasons of clarity the junctions on the operational level are omitted [according to 30].

for the processing of the stationary environment [e.g. 31–33].

The main part of the *tactical* level is the so-called context or scene modeling. An overview is given in Ulbrich et al. [52]. The focus is to put the independently perceived environmental features into an associative context. In a first step the "scenery" (stationary environment) is built up, e.g. by associating traffic lights and stationary obstacles to a certain lane. In a second step this scenery has to be combined with the movable environmental features to get a complete scene (according to the definitions in Geyer et al. [16]). In this step for example other traffic participants are associated to a certain lane (see e.g. Knaup and Homeier [26]). Hence, on this level of abstraction, the *semantic* information is the most relevant one besides the geometric and topologic information. Additionally, an abstracted vehicle state is also part of a complete scene which is transmitted to the mission accomplishment.

On the *strategical* level the environment is observed in an even more abstract way on a macro-scale level. The relevant features consist of the road network as well as information about the macroscopic traffic flow. The main information on this abstraction level is the *topology*, the connection of the roads, which is necessary for route planning. The geometric and semantic information is still available and necessary e.g. for planning optimized routes and following the planned routes.

4.4.2 Input and output

The data to be processed is received from three different sources (see Figure 3):

1. vehicle sensors (internal sensors according to Knoll and Christaller [27]),
2. environmental sensors (external sensors according to Knoll and Christaller [27])
3. external data (see Section 4.3)

In addition to classic architectures for autonomous vehicles, data derived from the environmental perception is no longer transmitted only to the mission accomplishment, but also to external receivers in a sense of a cooperative or collaborative mechanism (e.g. traffic participants in the close surrounding via V2V or central services via V2I communication, see Section 4.3). The acquisition of such data is already made today by mobile phone tracking or even manually: Drivers tell radio stations about current traffic jams or traffic controls and thus share their information with other traffic participants. This information is usually communicated with a map-relative position (example: on the motorway ... between junction A and junction B, 15 km jam due to an accident, 40 min delay).

4.4.3 Data processing

The perception module, covering environmental and self-perception, is the central module of an autonomous vehicle referring to this architecture proposition. All available information about the vehicle's surroundings and the vehicle's state are aggregated in this module and are prepared for a subsequent processing in the mission accomplishment.

At each level an abstraction of the data from the underlying level is required by introducing model assumptions. It is also conceptually possible to integrate data from external sources. Figure 3 gives an overview of the methods and processing steps of the environmental perception corresponding to their level of abstraction.

In addition to the environmental perception, the self-perception is also part of this column. Only detailed knowledge about the vehicle's state allows the autonomous vehicle to avoid decisions which might be hazardous for the system or the environment. The own state includes the states of each sensor and actor, fuel and battery levels as well as steering angle or rotation rates of the wheels.

A more detailed discussion of the components self-representation and self-perception is given in Maurer [34], Siedersberger [49], Pellkofer [44], or Bergmiller [5].

4.5 Mission accomplishment

4.5.1 Processed data

The mission accomplishment processes the defined mission. The processing is subdivided into three steps: planning at the strategical level, deciding at the tactical level and executing at the operational level.

According to the concepts of Donges [12] and Du et al. [14] the aforementioned different features are provided for each level of abstraction (see Figures 2 and 3):

- strategical level: road network and traffic flow, *macro-scale* information,
- tactical level: abstracted local scene, consisting of the scenery and the movable environmental features, *meso-scale* information,
- operational level: exact geometric values for a reactive collision avoidance and vehicle stabilization, *micro-scale* information.

4.5.2 Input and output

The way, a passenger can instruct an autonomous vehicle, is limited to the *strategical* level (according to Section 2). For a driver assistant system, additional possibilities of interaction between the vehicle and the passenger or driver exist at the other levels. For example, at the *tactical* level the driver may choose a maneuver or in the case of ACC a time gap to the vehicle ahead.

The results of the mission accomplishment do not only need to be executed (information flow downwards in Figure 1) but also have to be communicated (information flow to the right in Figure 1). This communication can be performed over multiple channels (optic, haptic, acoustic).

The receivers inside the vehicle are the passengers in the case of an automated vehicle. In case of assistant systems the receiver is the driver. Outside the vehicle, communication is directed to other traffic participants, and in a broader sense, also to animals. Communication is performed optically (e.g. turn indicators at the tactical level or the brake lights at the operational level), acoustically (e.g. the horn on tactical level) or via V2V-communication.

4.5.3 Data processing

Within the mission accomplishment the mission is concretized stepwise towards manipulating values taking environmental information into account. The process is illustrated in Figure 3.

The first task for accomplishing the mission is the navigation task, performed at the highest, the *strategical* level, according to the discussions in Donges [12]. The required data from the environment for an on-road vehicle is the road network with information about the current traffic flow.

Based on this data a route is planned which considers the optimization criteria defined by the passenger. Usually, this is done once per mission input. But due to changing information about the traffic flow or additionally detected roads at runtime the planned route can be adapted online. The resulting route should be communicated (visualized) to the passenger for verification purposes.

After the route planning is done, the environment changes relative to the vehicle due to the vehicle's movement. The output to the vehicle guidance at the tactical level is the next navigation point, already known from today's navigation systems (e.g. "In 500 m turn left"). In case of a driver assistance system this announcement is also an acoustic way of communication to the driver which can be complemented e.g. by current traffic disruptions.

The *tactical* level (or according to Donges [12] the guidance level) receives the mission indirectly by the input of the next navigation point from the strategical level. The perception provides the abstract and application-independent scene, including the vehicle's pose within that scene. The situation assessment now analyses the scene with respect to the current mission (next navigation point respectively) and extracts the relevant elements from the scene. If critical situations are detected, this result can be communicated to other traffic participants in different ways, e.g.:
- acoustically by the horn,
- optically using the headlight flasher or hazard lights, or
- via V2X to other automated vehicles

In case of an assistant system information or warnings in critical situations can also be communicated to the driver.

The decision unit selects the driving maneuvers based on the current situation with respect to the traffic regulations. In some cases, the selected maneuvers have to be communicated to the local environment (e.g. in the case of a lane change or a turning maneuver). This communication is also performed with the corresponding technical components as turn indicators or via V2X. The horn, turn indicators, back-up light, lower- and upper-beam headlights, as well as warnings in dangerous situations are the communication components at the *tactical* level. At this level a simple open-loop control is applied (according to Donges [12]).

The desired driving maneuver with the characteristic parameters is executed at the *operational* level. The environmental perception provides the features from the model-based filtering mechanisms of the sensor data. The trajectory generation (explicitly not trajectory *planning*, because the planning is typically done on the strategical level with a larger time horizon) calculates a time- and space-based nominal position of the vehicle and thus applies a closed-loop control of the vehicle considering the current environmental data. That is why the trajectory generation is classified as an overlaid closed-loop control.

The task of the subordinate closed-loop controller is the realization of the nominal positions generated by the trajectory calculation. The outputs of such subordinate closed-loop controllers are manipulated values for the engine, brakes or the steering system. In this way, this level is characterized by a direct feedback from internal and external sensors and thus performs a closed-loop control (argumentation similar to Donges [12]).

The brake lights during a braking maneuver or the trigger of an optic, haptic, or acoustic warning of the driver (e.g. in the case of an undesired lane departure) are parts of the communication options on the operational level from the viewpoint of the system.

5 Check against the functional requirements

According to Hertzberg et al. [20] the proposed architecture is a hybrid architecture combining the advantages of a sequential and a parallel architecture. Our approach can be understood as an extension of the work of Dickmanns [9], Maurer [34], and Pellkofer [44].

In our architecture, the aforementioned requirements for an autonomous on-road motor vehicle are considered as follows:
1. Operating: This architecture provides a bidirectional interface to the passenger (see Section 4.5).
2. Mission accomplishment: This architecture considers a widely used three-level approach to accomplish the mission: planning at the strategical level, deciding at the tactical level and executing at the operational level (see Section 4.5).
3. Map data: Map data is taken into account at several levels of abstraction within the system. Additionally, a transition of map data into adjacent levels of abstraction is described. Furthermore, mechanisms for automated map updates as well as the exchange of data describing the movable environment (traffic participants) is part of the concept (see Section 4.3).
4. Localization: The architecture considers two variants of global localization – absolute global localization which is only based on localization and motion sensors (see Section 4.2) and map-relative localization (see Section 4.3). Furthermore, a local position estimation based on motion estimation is part of the concept (see Section 4.4). In case that a global position is

not part of the system, approaches solving the SLAM-problem are considered as a local map-relative pose by this architecture (see Section 4.3). Thus, we have incorporated four different localization solutions (in extension to Moore et al. [39]) .

5. Environmental perception: The central element of the architecture is the environmental perception. In this module all external information is aggregated to a consistent image of the vehicle's surroundings (see Section 4.4).
6. Cooperation: Cooperative approaches are considered in multiple respects. The concept allows the vehicle to communicate environmental data (collaboration), intentions or manipulated values to its environment (see Section 4.4 and 4.5). This information can also be received as external data from other traffic participants (see Section 4.3). Cooperation in the sense of an explicit interaction with other traffic participants is mainly located in the module "guidance" on the tactical level.
7. Safety: This aspect is partly covered by the self-perception and its integration into the scene. Furthermore, the architecture makes functional redundancy transparent: theoretically, required environmental data can be derived completely from external data or completely from the environmental perception or – and this might be a case of parallel redundancy – from both external data and environmental perception. If information of these two independant sources does not match any longer, corresponding mechanisms have to be applied to remain in a safe state, because in case of autonomous driving the driver is no longer available as a fallback solution.
8. Self-perception: As a part of the perception the self-perception provides information about the state of the vehicle and its motion (see Section 4.4).

Moreover, today's driver assistant systems can be assigned to the introduced abstraction levels. Systems such as ESC, anti-lock braking system (ABS), lane departure warning (LDW) etc. mainly operate at the *operational* level. Their tasks focus on the stabilization of the vehicle, concerning the physical limits in the case of an ABS or ESC or concerning the environment (LDW) with a strongly limited preview. An ACC-system already works at both, the *operational* and the *tactical* level. In particular, the selection of the relevant ACC-target is a process of the tactical level. On the contrary, the closed-loop control of the distance to the vehicle ahead is part of the operational level. Today's navigation systems are mainly located on the strategical level.

6 Conclusion

In this article we have proposed a holistic functional system architecture for autonomous on-road motor vehicles which extends existing architectures by a systematic integration of external data such as map data and V2X-information. We explicitly combined perception-driven (e.g. Leonard et al. [28]) and localization-driven (e.g. Bacha et al. [2], Montemerlo et al. [38], Nothdurft et al. [43], Rauskolb et al. [47], Wille et al. [56]) approaches for autonomous driving in one single system description. It thus points out two different ways of information flow (see Figure 1): one short loop directly from the external sensors through the environmental perception to the mission accomplishment and a second larger loop over a global localization and external data to the mission accomplishment. Additionally, the consideration of a bidirectional communication allows the implementation of automated map updates and we have identified four different localization solutions.

The proposed system architecture also allows the discussion about the role of the tactical level in future systems. Advanced driver assistant systems, like an intersection assistant (see e.g. Mages et al. [29] or Herrmann [19]), make it clear that larger time horizons are required for collision free driving. The prediction only based on kinematic values does not yield usable results, especially in inner-city scenarios. In these cases it is necessary to consider at least the scenery (e.g. lane course and traffic lights) or even the scene (including other traffic participants) for the prediction of the movable environment. Furthermore, cooperative driving and the consideration of local traffic regulations are mainly performed on the tactical level. That is why we expect the tactical level to take on an added importance in future systems.

Acknowledgement: The authors like to thank all active members and alumni of the research project Stadtpilot, who all contributed by their work and countless discussions to this article.

References

1. Albus, J. S. (2002). 4d/RCS: a reference model architecture for intelligent unmanned ground vehicles. In *AeroSense 2002*.
2. Bacha, A., Bauman, C., Faruque, R., Fleming, M., Terwelp, C., Reinholtz, C., Hong, D., Wicks, A., Alberi, T., and Anderson, D. (2008). Odin: Team VictorTango's entry in the DARPA urban challenge. *Journal of Field Robotics*, 25(8):467–492.
3. Baker, C. R. and Dolan, J. M. (2008). Traffic interaction in the urban challenge: Putting boss on its best behavior. In *IEEE/RSJ*

International Conference on Intelligent Robots and Systems (IROS).
4. Bar Hillel, A., Lerner, R., Levi, D., and Raz, G. (2012). Recent progress in road and lane detection: a survey. *Machine Vision and Applications*.
5. Bergmiller, P. (2014). *Towards Functional Safety in Drive-by-Wire Vehicles*. draft version phd thesis, Technische Universität Braunschweig, Braunschweig.
6. Bohren, J., Foote, T., Keller, J., Kushleyev, A., Lee, D., Stewart, A., Vernaza, P., Derenick, J., Spletzer, J., and Satterfield, B. (2008). Little ben: The ben franklin racing team's entry in the 2007 DARPA urban challenge. *Journal of Field Robotics*, 25(9):598–614.
7. Bonasso, P., Firby, J., Gat, E., Kortenkamp, D., Miller, D. P., and Slack, M. G. (1997). Experiences with an architecture for intelligent, reactive agents. *Journal of Experimental & Theoretical Artificial Intelligence*, 9(2-3):237–256.
8. Deragården, H. and Thomas, B. (2007). Specification of the map data feedback loop test environment. RE-Report 4.1, Ertico.
9. Dickmanns, E. D. (2007). *Dynamic Vision for Perception and Control of Motion*. Springer London.
10. Dickmanns, E. D., Behringer, R., Dickmanns, D., Hildebrandt, T., Maurer, M., Thomanek, F., and Schiehlen, J. (1994). The seeing passenger car'VaMoRs-p'. In *IEEE Intelligent Vehicles Symposium (IV)*.
11. Donges, E. (1982). Aspekte der aktiven Sicherheit bei der Führung von Personenkraftwagen. *Automobil-Industrie*, 27:183–190.
12. Donges, E. (1999). A conceptual framework for active safety in road traffic. *Vehicle System Dynamics*, 32(2-3):113–128.
13. Donges, E. (2012). Fahrerverhaltensmodelle. In Winner, H., Hakuli, S., and Wolf, G., editors, *Handbuch Fahrerassistenzsysteme: Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort*, ATZ-MTZ-Fachbuch, pages 15–23. Vieweg + Teubner, 2nd edition.
14. Du, J., Masters, J., and Barth, M. (2004). Lane-level positioning for in-vehicle navigation and automated vehicle location (AVL) systems. In *7th International IEEE Conference on Intelligent Transportation Systems (ITSC)*.
15. Elfes, A. (1987). Sonar-based real-world mapping and navigation. *IEEE Journal of Robotics and Automation*, 3(3):249–265.
16. Geyer, S., Baltzer, M., Franz, B., Hakuli, S., Kauer, M., Kienle, M., Meier, S., Weißgerber, T., Bengler, K., Bruder, R., Flemisch, F., and Winner, H. (2014). Concept and development of a unified ontology for generating test and use-case catalogues for assisted and automated vehicle guidance. *IET Intelligent Transport Systems*, 8(3):183–189.
17. Hale, A. R., Stoop, J., and Hommels, J. (1990). Human error models as predictors of accident scenarios for designers in road transport systems. *Ergonomics*, 33(10-11):1377–1387.
18. Hamilton-Baillie, B. (2008). Shared space: reconciling people, places and traffic. *Built environment*, 34(2):161–181.
19. Herrmann, S. (2013). *Kollisionswarnung im urbanen Straßenverkehr auf Basis einer probabilistischen Situationsanalyse*. PhD thesis, Technische Universität Braunschweig, Braunschweig.
20. Hertzberg, J., Lingemann, K., and Nüchter, A. (2012). Roboterkontrollarchitekturen. In *Mobile Roboter*, pages 317–333. Springer Berlin Heidelberg, Berlin, Heidelberg.
21. Holder, S., Hörwick, M., and Gentner, H. (2012). Funktionsübergreifende Szeneninterpretation zur Vernetzung von Fahrerassistenzsystemen. In *Automatisierungssysteme, Assistenzsysteme und eingebettete Systeme für Transportmittel (AAET)*.
22. Homeier, K. and Wolf, L. (2011). RoadGraph: high level sensor data fusion between objects and street network. In *14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*.
23. Huang, A. S., Moore, D., Antone, M., Olson, E., and Teller, S. (2009). Finding multiple lanes in urban road networks with vision and lidar. *Autonomous Robots*, 26(2-3):103–122.
24. Hurdus, J., Bacha, A., Bauman, C., Cacciola, S., Faruque, R., King, P., Terwelp, C., Currier, P., Hong, D., Wicks, A., and Reinholtz, C. (2008). VictorTango architecture for autonomous navigation in the DARPA urban challenge. *Journal of Aerospace Computing, Information, and Communication*, 5(12):506–529.
25. Kammel, S., Ziegler, J., Pitzer, B., Werling, M., Gindele, T., Jagzent, D., Schröder, J., Thuy, M., Goebl, M., Hundelshausen, F. v., Pink, O., Frese, C., and Stiller, C. (2008). Team AnnieWAY's autonomous system for the 2007 DARPA urban challenge. *Journal of Field Robotics*, 25(9):615–639.
26. Knaup, J. and Homeier, K. (2010). RoadGraph - graph based environmental modelling and function independent situation analysis for driver assistance systems. In *13th International IEEE Conference on Intelligent Transportation Systems (ITSC)*.
27. Knoll, A. and Christaller, T. (2000). Selbstrepräsentation, Selbstwahrnehmung und Verhaltenssteuerung bei Robotern. In Sandkühler, H. J. and Stekeler-Weithofer, P., editors, *Philosophie und Geschichte der Wissenschaften*, pages 243–270. Verlag Peter Lang.
28. Leonard, J., How, J., Teller, S., Berger, M., Campbell, S., Fiore, G., Fletcher, L., Frazzoli, E., Huang, A., and Karaman, S. (2008). A perception-driven autonomous urban vehicle. *Journal of Field Robotics*, 25(10):727–774.
29. Mages, M., Hopstock, M., and Klanner, F. (2012). Kreuzungsassistenz. In Winner, H., Hakuli, S., and Wolf, G., editors, *Handbuch Fahrerassistenzsysteme: Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort*, ATZ-MTZ-Fachbuch, pages 572–581. Vieweg + Teubner, 2nd edition.
30. Matthaei, R. (2014). *Wahrnehmungsgestützte Lokalisierung in fahrstreifengenauen Karten für Assistenzsysteme und automatisches Fahren in urbaner Umgebung*. draft version phd thesis, Technische Universität Braunschweig, Braunschweig.
31. Matthaei, R., Bagschik, G., and Maurer, M. (2014a). Map-relative localization in lane-level maps for ADAS and autonomous driving. In *IEEE Intelligent Vehicles Symposium (IV)*.
32. Matthaei, R., Bagschik, G., Rieken, J., and Maurer, M. (2014b). Stationary urban environment modeling using multi-layer-grids. In *Proceedings of the 17th International Conference on Information Fusion (FUSION)*.
33. Matthaei, R., Lichte, B., and Maurer, M. (2013). Robust grid-based road detection for ADAS and autonomous vehicles in urban environments. In *Proceedings of the 16th International Conference on Information Fusion (FUSION)*.
34. Maurer, M. (2000). *Flexible Automatisierung von Straßenfahrzeugen mit Rechnersehen*. PhD thesis, Universität der Bundeswehr München, Neubiberg.
35. Michon, J. A. (1985). *A critical view of driver behavior models: what do we know, what should we do?* Springer.

36. Milford, M. and Wyeth, G. (2008). Mapping a suburb with a single camera using a biologically inspired SLAM system. *IEEE Transactions on Robotics*, 24(5):1038–1053.
37. Miller, I., Campbell, M., Huttenlocher, D., Kline, F.-R., Nathan, A., Lupashin, S., Catlin, J., Schimpf, B., Moran, P., Zych, N., Garcia, E., Kurdziel, M., and Fujishima, H. (2008). Team cornell's skynet: Robust perception and planning in an urban environment. *Journal of Field Robotics*, 25(8):493–527.
38. Montemerlo, M., Becker, J., Bhat, S., Dahlkamp, H., Dolgov, D., Ettinger, S., Haehnel, D., Hilden, T., Hoffmann, G., Huhnke, B., Johnston, D., Klumpp, S., Langer, D., Levandowski, A., Levinson, J., Marcil, J., Orenstein, D., Paefgen, J., Penny, I., Petrovskaya, A., Pflueger, M., Stanek, G., Stavens, D., Vogt, A., and Thrun, S. (2008). Junior: The stanford entry in the urban challenge. *Journal of Field Robotics*, 25(9):569–597.
39. Moore, D. C., Huang, A. S., Walter, M., Olson, E., Fletcher, L., Leonard, J., and Teller, S. (2009). Simultaneous local and global state estimation for robotic navigation. In *IEEE International Conference on Robotics and Automation (ICRA)*.
40. Müller, A., Himmelsbach, M., Lüttel, T., Hundelshausen, F. v., and Wünsche, H.-J. (2011). GIS-based topological robot localization through LIDAR crossroad detection. In *14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*.
41. Nilsson, J.-O., Zachariah, D., and Skog, I. (2012). Global navigation satellite systems: An enabler for in-vehicle navigation. In Eskandarian, A., editor, *Handbook of Intelligent Vehicles*, pages 311–342. Springer London, London.
42. Nothdurft, T. (2014). *Ein Kontextmodell für sicherheitsrelevante Anwendungen in der autonomen Fahrzeugführung*. PhD thesis, Technische Universität Braunschweig, Braunschweig.
43. Nothdurft, T., Hecker, P., Ohl, S., Saust, F., Maurer, M., Reschka, A., and Bohmer, J. R. (2011). Stadtpilot: First fully autonomous test drives in urban traffic. In *14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*.
44. Pellkofer, M. (2003). *Verhaltensentscheidung für autonome Fahrzeuge mit Blickrichtungssteuerung*. PhD thesis, Universität der Bundeswehr München, Neubiberg.
45. Quddus, M. A., Ochieng, W. Y., and Noland, R. B. (2007). Current map-matching algorithms for transport applications: State-of-the art and future research directions. *Transportation Research Part C: Emerging Technologies*, 15(5):312–328.
46. Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, 13(3):257–266.
47. Rauskolb, F. W., Berger, K., Lipski, C., Magnor, M., Cornelsen, K., Effertz, J., Form, T., Graefe, F., Ohl, S., Schumacher, W., Wille, J.-M., Hecker, P., Nothdurft, T., Doering, M., Homeier, K., Morgenroth, J., Wolf, L., Basarke, C., Berger, C., Gülke, T., Klose, F., and Rumpe, B. (2008). Caroline: An autonomously driving vehicle for urban environments. *Journal of Field Robotics*, 25(9):674–724.
48. SAE international (2014). Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems.
49. Siedersberger, K.-H. (2003). *Komponenten zur automatischen Fahrzeugführung in sehenden (semi-)autonomen Fahrzeugen*. PhD thesis, Universität der Bundeswehr München, Neubiberg.
50. Szczot, M., Serfling, M., Löhlein, O., Schüle, F., Konrad, M., and Dietmayer, K. (2010). Global positioning using a digital map and an imaging radar sensor. In *IEEE Intelligent Vehicles Symposium (IV)*.
51. Thrun, S., Burgard, W., and Fox, D. (2005). *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents series)*. Intelligent robotics and autonomous agents. The MIT Press.
52. Ulbrich, S., Nothdurft, T., Maurer, M., and Hecker, P. (2014). Graph-based context representation, environment modeling and information aggregation for automated driving. In *IEEE Intelligent Vehicles Symposium (IV)*.
53. Visintainer, F. and Darin, M. (2008). Final requirements and strategies for map feedback. Report D2.2, Ertico.
54. Wachenfeld, W. and Winner, H. (2013). Use Cases des Autonomen Fahrens. Technical Report Release 1, Villa Ladenburg Kolleg „Autonomes Fahren“.
55. Wille, J. M. (2012). *Manöverübergreifende autonome Fahrzeugführung in innerstädtischen Szenarien am Beispiel des Stadtpilotprojekts*. PhD thesis, Technische Universität Braunschweig, Braunschweig.
56. Wille, J. M., Saust, F., and Maurer, M. (2010). Stadtpilot: Driving autonomously on Braunschweig's inner ring road. In *IEEE Intelligent Vehicles Symposium (IV)*.

Bionotes

Richard Matthaei
Technische Universität Braunschweig, Institut für Regelungstechnik, 38106 Braunschweig
matthaei@ifr.ing.tu-bs.de

Richard Matthaei is scientific assistant with the Institute of Control Engineering at the Technische Universität Braunschweig. Main fields of research: functional system architecture for autonomous vehicles, environmental perception and map-relative localization in urban environments.

Prof. Dr.-Ing. Markus Maurer
Technische Universität Braunschweig, Institut für Regelungstechnik, 38106 Braunschweig
maurer@ifr.ing.tu-bs.de

Prof. Dr.-Ing. Markus Maurer is director of the department Vehicle Electronics of the Institute of Control Engineering at the Technische Universität Braunschweig. Main fields of research: autonomous driving/driver assistant systems, electromobility and vehicle systems engineering.